Opponent-Shaping and Interference in General-Sum Games

Dr Jakob Foerster

University of Oxford

Monday 23 February 2023
2.00pm
MR4, IEB

Abstract:

In general-sum games, the interaction of self-interested learning agents commonly leads to collectively worst-case outcomes, such as defect-defect in the iterated prisoner's dilemma (IPD). To overcome this, some methods, such as Learning with Opponent-Learning Awareness (LOLA), shape their opponents' learning process. However, these methods are myopic since only a small number of steps can be anticipated, are asymmetric since they treat other agents as naive learners, and require the use of higher-order derivatives, which are calculated through white-box access to an opponent's differentiable learning algorithm.

In this talk I will first introduce Model-Free Opponent Shaping (M-FOS), which overcomes all of these limitations. M-FOS learns in a meta-game in which each meta-step is an episode of the underlying (``inner'') game. The meta-state consists of the inner policies, and the meta-policy produces a new inner policy to be used in the next episode. M-FOS then uses generic model-free optimisation methods to learn meta-policies that accomplish long-horizon opponent shaping.

I will finish off the talk with our recent results for adversarial (or cooperative) cheap-talk: How can agents interfere with (or support) the learning process of other agents without being able to act in the environment?